

WHAT IS CLAIMED IS:

1. A method for processing a data string of Arabic text characters into Arabic calligraphic script representation data, the method comprising:

identifying words in said string;

identifying a form of said characters in said words, said form comprising initial, medial, final and isolated;

for said characters that are not of said isolated form, identifying a type of said characters as a function of compatibility with a type of a neighboring character;

selecting, for each one of said characters in said data string, a glyph from a set of predetermined glyphs corresponding to said characters, said form and said type; and

determining a vertical offset for each said glyph to match neighboring glyphs, said script representation data comprising glyph identification data and offset data for each character in said data string.

2. The method as claimed in claim 1, wherein said type identified is a best match of attributes between glyphs available in said set of glyphs for said form of said characters, said best match corresponding to a visualization of a calligrapher.

3. The method as claimed in claim 2, wherein said attributes are thickness, pen movement direction, pen rotation direction, and waveform.

4. The method as claimed in any one of claims 1 to 3, wherein said set of glyphs comprises 512 glyphs or fewer.

5. The method as claimed in any one of claims 1 to 3, wherein said set of glyphs comprises 256 glyphs or fewer.

6. The method as claimed in any one of claims 1 to 5, wherein diacritics are represented by separate characters in said string corresponding to separate glyphs in said set of predetermined glyphs, said selecting comprising determining

an offset position of each diacritic to be associated with a glyph representing a letter.

7. The method as claimed in claim 6, wherein unacceptable combination of diacritics is verified and disallowed.

8. The method as claimed in any one of claims 1 to 7, wherein some of said glyphs represent ligatures.

9. The method as claimed in claim 8, wherein said ligatures represent ligatures joining two letters.

10. The method as claimed in claim 8, wherein said ligatures represent ligatures joining a letter and at least one diacritic.

11. An apparatus for processing a data string of Arabic text characters output from an Arabic text source into Arabic calligraphic script representation data, the apparatus comprising:

a word identification module receiving said data string and outputting a word;

a form identification module receiving said word and outputting a form of said characters in said word, said form being one of initial, medial, final, and isolated;

a type identification module receiving said form and said characters and outputting type data of said characters as a function of compatibility with a type of a neighboring character;

a glyph identification module receiving said type data and said characters and selecting, for each one of said characters, a glyph from a set of predetermined glyphs corresponding to said characters, said form, and said type; and

an offset determining module receiving said glyph and said characters and determining a vertical offset for said glyph to match neighboring glyphs and outputting said calligraphic script representation data.

12. An apparatus as claimed in claim 11, wherein said type identification module identifies a best match of attributes between glyphs available in a set of glyphs for a form of a character, said best match corresponding to a visualization of a calligrapher.
13. An apparatus as claimed in claim 12, wherein said attributes are thickness, pen movement direction, pen rotation direction, and waveform.
14. An apparatus as claimed in any one of claims 11 to 13, wherein said set of glyphs comprises 512 glyphs or fewer.
15. An apparatus as claimed in any one of claims 11 to 13, wherein said set of glyphs comprises 256 glyphs or fewer.
16. An apparatus as claimed in any one of claims 11 to 15, wherein said word identification module identifies diacritics as separate characters in said string, said glyph identification module associates said diacritics to separate glyphs in said set of predetermined glyphs, and said offset determining module determines an offset position of each diacritic to be associated with a glyph representing a letter.
17. An apparatus as claimed in claim 16, wherein said word identification module verifies unacceptable combinations of diacritics disallows them.
18. An apparatus as claimed in any one of claims 11 to 17, wherein some of said glyphs in said predetermined set of glyphs represent ligatures.
19. An apparatus as claimed in claim 18, wherein said ligatures represent ligatures joining two letters.
20. An electronic printing apparatus comprising an apparatus for processing a data string of Arabic text characters output from an Arabic text source into Arabic calligraphic script representation data, said apparatus for processing defined as

per any one of claims 11 to 19.

21. An electronic printing apparatus as claimed in claim 20, wherein said electronic printing apparatus comprises an input/output module, and said apparatus for processing inputs said calligraphic script representation data into said input/output module.

22. An electronic printing apparatus as claimed in claim 20, wherein said electronic printing apparatus comprises an image/text translator, and said apparatus for processing inputs said calligraphic script representation data into said image/text translator.

23. A user device having an application and a printer device driver and comprising an apparatus for processing a data string of Arabic text characters output from an Arabic text source into Arabic calligraphic script representation data, said apparatus for processing defined as per any one of claims 11 to 19.

24. A web-browser comprising a translator that takes standard text in HTML and converts it onscreen, characterized in that the translator comprises an apparatus for processing a data string of Arabic text characters output from an Arabic text source into Arabic calligraphic script representation data, said apparatus for processing defined as per any one of claims 11 to 19.